



PTO/SB/21 (03-03)

Approved for use through 04/30/2003. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Total Number of Pages in This Submission	Application Number	10/761,885
	Filing Date	January 21, 2004
	First Named Inventor	Hitoshi Fujimiya
	Art Unit	Not known
	Examiner Name	Not known
	Attorney Docket Number	Y03S013-US

ENCLOSURES (Check all that apply)		
<input type="checkbox"/> Fee Transmittal Form <input type="checkbox"/> Fee Attached <input type="checkbox"/> Amendment/Reply <input type="checkbox"/> After Final <input type="checkbox"/> Affidavits/declaration(s) <input type="checkbox"/> Extension of Time Request <input type="checkbox"/> Express Abandonment Request <input type="checkbox"/> Information Disclosure Statement <input checked="" type="checkbox"/> Certified Copy of Priority Document(s) <input type="checkbox"/> Response to Missing Parts/ Incomplete Application <input type="checkbox"/> Response to Missing Parts under 37 CFR 1.52 or 1.53	<input type="checkbox"/> Drawing(s) <input type="checkbox"/> Licensing-related Papers <input type="checkbox"/> Petition <input type="checkbox"/> Petition to Convert to a Provisional Application <input type="checkbox"/> Power of Attorney, Revocation <input type="checkbox"/> Change of Correspondence Address <input type="checkbox"/> Terminal Disclaimer <input type="checkbox"/> Request for Refund <input type="checkbox"/> CD, Number of CD(s) _____ <div style="border: 1px solid black; padding: 2px; margin-top: 5px;">Remarks</div>	<input type="checkbox"/> After Allowance Communication to a Technology Center (TC) <input type="checkbox"/> Appeal Communication to Board of Appeals and Interferences <input type="checkbox"/> Appeal Communication to TC (Appeal Notice, Brief, Reply Brief) <input type="checkbox"/> Proprietary Information <input type="checkbox"/> Status Letter <input type="checkbox"/> Other Enclosure(s) (please identify below):

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm or Individual	Konomi Takeshita
Signature	
Date	March 12, 2004

CERTIFICATE OF TRANSMISSION/MAILING

I hereby certify that this correspondence is being facsimile transmitted to the USPTO or deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Commissioner for Patents, Washington, DC 20231 on this date: _____

Typed or printed	Kurara Yamagishi	
Signature		Date 3/15/04

This collection of information is required by 37 CFR 1.5. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, Washington, DC 20231.

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2003年 1月21日
Date of Application:

出願番号 特願2003-048216
Application Number:

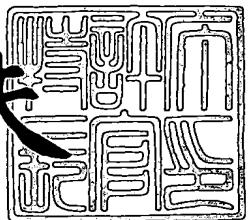
[ST. 10/C] : [JP2003-048216]

出願人 株式会社ダイナコム
Applicant(s):

2004年 2月 9日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫



【書類名】 特許願
【整理番号】 DCP2003011
【提出日】 平成15年 1月21日
【あて先】 特許庁長官 殿
【発明の名称】 遺伝子連鎖不平衡の表示方法
【請求項の数】 1
【発明者】
【住所又は居所】 千葉県茂原市茂原 643 株式会社ダイナコム内
【氏名】 中村 英二
【発明者】
【住所又は居所】 千葉県茂原市茂原 643 株式会社ダイナコム内
【氏名】 安達 宏紀
【発明者】
【住所又は居所】 千葉県茂原市茂原 643 株式会社ダイナコム内
【氏名】 藤宮 仁
【特許出願人】
【識別番号】 597101155
【住所又は居所】 千葉県茂原市茂原 643
【氏名又は名称】 株式会社ダイナコム
【代表者】 藤宮 仁

【書類名】 明細書

【発明の名称】 遺伝子連鎖不平衡の表示方法

【特許請求の範囲】

【請求項 1】 2つ以上の遺伝子座位における遺伝子多様性データを入力するためのデータ入力手段と、前記手段で入力したデータを記憶するためのデータ記憶手段と、前記データ記憶手段から取り出した任意の2つの遺伝子座位の連鎖不平衡値を計算する連鎖不平衡値算出手段と、前記連鎖不平衡値算出手段から得られた結果を表示するための連鎖不平衡値表示手段とを備えた連鎖不平衡の計算システムであって、第一の遺伝子多様性データ群から得られた各遺伝子座位の相互の遺伝的関係の度合いと、第二の遺伝子多様性データ群から得られた各遺伝子座位の連鎖不平衡値をそれぞれ別の色に対応付けて混合するか、または第一および第二の連鎖不平衡値の差を求め、関連する色をつけて表示することを特徴とする連鎖不平衡結果の表示方法

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は遺伝子多様性データ解析に関し、ケースデータ群とコントロールデータ群のそれぞれで求められたペアワイズ連鎖不平衡値を表示する上で、ケースデータの処理結果と、コントロールデータ群での処理結果を見やすく比較表示するための方法に関する。

【0002】

【従来の技術】 遺伝子多様性研究では、各遺伝子座位相互の連鎖の強さを計算することが頻繁に行われる。連鎖とはある遺伝子座位の多型と別に着目する座位の遺伝子多型がペアで子孫に遺伝していることを意味している。もし染色体上で十分離れていれば、遺伝子のランダムな組み換えが起こるため、5, 6世代経過後にはほぼ平衡状態に落ち着くことが知られている。この状態をハーディワイバーグ平衡と呼ぶ。注目する遺伝子多様性の座位が物理的に近い場合、このハーディワインバーグ平衡からずれが保存される。このずれを連鎖不平衡と呼ぶ。

【0003】

連鎖不平衡は2箇所のハプロタイプ度数情報を用いて、 2×2 の分割表を作成し、各座位でのハプロタイプ頻度から想定される独立の場合からどれだけずれているかを連鎖不平衡値として用いる。

まず、第一の遺伝子座位と第二の遺伝子座位のメジャーアレルを1、マイナーアレルを3としてそれぞれのハプロタイプ頻度を以下のように表す。

第一遺伝子座位・第二遺伝子座位 頻度

1 - 1	p 1 1
1 - 3	p 1 3
3 - 1	p 3 1
3 - 3	p 3 3

すると、連鎖不平衡Dは次式で与えられる。

$$D = p_{11}p_{33} - p_{13}p_{31}$$

【0004】

Dは正負の値をとるが、0 - 1間の値をとるように補正したD' という連鎖不平衡値も定義されている。D' は $D > 0$ or $D = 0$ の場合は、Dの取り得る最大値は、次式で与えられる。

$$D_{\max} = \min(p_{1\triangle} \times p_{\triangle 3}, p_{3\triangle} \times p_{\triangle 1})$$

ただし、 $p_{1\triangle}$ は第一座位のメジャーアレル頻度、 $p_{\triangle 3}$ は第二座位のマイナーアレル頻度、同様に $p_{3\triangle}$ は第一座位のマイナーアレル頻度、 $p_{\triangle 1}$ は第二座位のメジャーアレル頻度を意味する。

$D < 0$ の場合はDの取り得る最小値は次式で与えられる。

$$D_{\min} = \max(-p_{1\triangle} \times p_{\triangle 1}, -p_{3\triangle} \times p_{\triangle 3})$$

これらを用い、

$$D' = D / D_{\max} \quad (D \text{が正の場合})$$

$$D' = D / D_{\min} \quad (D \text{が負の場合})$$

と定義される。

【0005】

また、他に r^2 と呼ばれる連鎖不平衡値があり、次式で表される。

$$r^2 = D^2 / (p_{1\triangle} \times p_{3\triangle} \times p_{\triangle 1} \times p_{\triangle 3})$$

そのほかにも赤池情報量基準（以下A I Cと呼ぶ。Akaike's Information Criterionの略）を用いた方法などがある（K. Shimono-noda et al: Akaike's information criterion for a measure of linkage disequilibrium, Journal of Human Genetics, Vol. 47 Issue 12 (2002) pp 649-655）。

【0006】

これらの連鎖不平衡を表す指標値をケースデータ群とコントロールデータ群に対して求めることで疾患などのケース特有の連鎖不平衡の違いを持つ部分を見つけることが可能となる。図2に一塩基多型（図中ではSNPと表示）の場合の入力データの例を示す。同図はヒトの2倍体の一塩基多型を検査した結果の例である。本データはメジャーアレルのホモを1に、マイナーアレルのホモを3に、メジャーアレルとマイナーアレルのヘテロを2としている。ここでメジャーアレルとは一般的にもっとも多い多型を意味する。また、マイナーアレルは対立遺伝子のひとつで、数の相対的に少ない多型を意味する。2倍体の検査結果であるから、メジャーアレルまたはマイナーアレルいずれか一方の2つを持っている場合はホモ、それぞれひとつが混在している場合はヘテロと呼ばれる。図中「群」というカラムは0がケース（罹患）、1がコントロール（健常者）を意味する。

【0007】

群のデータに基づきデータを2群にわけ、前記連鎖不平衡値 r^2 を算出したのが図2のケースデータ群の連鎖不平衡値とコントロールデータ群の連鎖不平衡値の一覧表である。各SNP座位のデータを前記式で説明したとおり、ペアワイズで分割表を作成し、それらの結果をまとめたものである。自分自身は連鎖不平衡が定義されないため、空欄である（完全に連鎖していると定義することも可能）。また、まったくの対象行列となるため、上三角行列だけを示してある。

【0008】

r^2 では0に近い値の場合は両座位にはあまり連鎖がないことを意味する。1に近い場合は強い連鎖を持つことを意味する。図2ではSNP1とSNP3が強い

連鎖をもち、SNP2とSNP4が強い連鎖を持っていることを示している。連鎖不平衡の計算結果としてケースデータの連鎖不平衡値とコントロールデータの連鎖不平衡値を書く対応するセルごとに比較することにより、両群の連鎖の度合いの違いがある部分を見つけることができる。たとえば図2ではSNP4の列が若干違った値をとっており、ケース群とコントロール群のデータに差があることを示している。

【0009】

【発明が解決しようとする課題】 しかしながら、従来の技術では、単に連鎖不平衡の指標値を表形式で別々に表示しているだけであった。したがって、ケース・コントロール間での相違箇所を見つけ出すのが大変であるという問題があった。また、一塩基多型の検査データは数十個から多くは数千個以上まで、対象とするため、全体的に見ながら相違点を見つけだすのが難しいという問題があった。

【0010】

【課題を解決するための手段】 本発明では上記のような課題を解決するために、各遺伝子座位のペアワイズでの連鎖不平衡を算出した後で、マトリクス状にケースとコントロールデータ群の連鎖不平衡値を配置し、おのおの別の色を与えて、混合して表示する手段を与える。また、ケース群から得られた連鎖不平衡とコントロール群から得られた連鎖不平衡値の差またはその絶対値を求め、その値に応じて色を変えながら表示する手段を与える。

【0011】

上記手段によれば、色の偏りや差分の大きさによる色の違いなどにより容易に2群での連鎖不平衡状態が異なっている部分を知ることができる。

【0012】

【発明の実施の形態】 以下、本発明の一実施例を図に示しながら説明する。図は発明を実施する形態の一例である。図1は、本発明の方法を説明するための全体構成図である。システムは遺伝子多様性検査データを入力するためのデータ入力手段10、そのデータを記憶するためのデータ記憶手段11、入力されたデータから群別にペアワイズの分割表を作成して連鎖不平衡値を算出する連鎖不

平衡値算出手段12、さらにその連鎖不平衡値を参照しながら色の濃度に対応させるための色算出手段13、算定された色データをマトリックス上に整列してグラフィック表示を行うための出力表示手段14から構成される。

【0013】

まず、従来の技術の項で説明した図2に示す入力データをデータ入力手段10で入力する。入力されたデータはデータ記憶手段11でメモリ内に保存される。さらに記憶したデータは必要に応じて補助記憶装置（図示せず）などにコピーして保存する。次に記憶されたデータを順次参照しながらペアワイズで各遺伝子座位の多様性データを参照し、「従来の技術」の項で説明したとおり、群のデータで「0」のケース群と「1」のコントロール群にデータを分類し、それぞれの群に対してペアワイズな全ての組み合わせで各遺伝子座位の分割表を作成する。

【0014】

分割表を元にD、D'、 r^2 、AICなどを指定条件の連鎖不平衡値算出手段12で算出する。さらに色算出手段13において与えられた変換式に基づいて色の濃さ」0から255までの値が割り当てられる。色の割り当てが決まつたら、出力表示手段14においてマトリックス状に並べて表示を行う。

【0015】

図4に色割付の処理フローを示す。まず、算出されたケース群の特定の座位間のペアワイズ連鎖不平衡値を参照する（40）。次にコントロール群に対して算出した同一ペアワイズの連鎖不平衡値を参照する（41）。次にあらかじめ設定してある表示方法に従って色の濃さを算出する。たとえば連鎖不平衡値をそのまま色に変換する場合は0.0～1.0間での値を0～255までのグラフィックス表示の濃さに割り当てる。差分で表示する場合はケース群の連鎖不平衡値とコントロール値を基準として連鎖不平衡値の差を求め、-1.0～0の負の値の場合は青色に、0～1.0の正の値には赤をそれぞれ絶対値にあわせて濃くするように割り当てる（42）。得られたセルの描画色のデータに基づき、画像に表示する（43）。次のセルの計算に移る（44）。さらにセルの終了かどうかを判定し、終了でなければ上に戻り、終了の場合はこの処理系を抜ける（45）。

【0016】

図3にグラフィック表示した例を示す。本実施例ではケース群のデータを緑に、コントロール群のデータを赤に割り付けている。したがってセル31は図2において該当するセルの値が両者とも0.1であり、緑、赤の色が薄く同レベルで混色され、黄色の薄い色となっている。セル32はケース、コントロールとも0.9であり、濃い黄色になっている。さらにセル33は、ケースが0.1で、コントロールが0.0であるため薄い緑である。セル34は、ケースが0.9でコントロールが1.0であり、濃い黄色であるがわずかに赤が強いため、オレンジにやや近い値となっている。

【0016】

このように、ケース、コントロール群の値をそのまま別の色に割り付けて重ねて混合して表示することにより、色の偏りがある場合はそこに連鎖不平衡の差があることを全体的に見ながら認識することができる。

【0017】

さらに、図3（B）の差分表示では、ケース群とコントロール群のデータの差を取り、差異のある場所のみを表示している。セル35は、コントロールを基準にしてケースが0.1だけ大きい場合である。大きい場合は赤に割り付けている。また、逆にコントロール群の値よりもケース群の値が小さい場合は青に割り付けている。すなわち、-1.0～0未満は青に、0～1.0は赤に割り当てる。また両者とも絶対値が大きいほど濃い色に割り当っている。差分表示では、両者の差がどの座位間に存在するか一目で知ることができる。

【0018】

このように本実施例によれば、ケース群とコントロール群の連鎖不平衡の違いを容易に見つけられるような表示方法が可能である。また、ケース・コントロールに限らず、別の特長による集計を適用して連鎖不平衡を求め、それらの違いを表示することも可能である。3つ以上の群を持つ場合はそれぞれ基準の群に対して差を求め、それぞれ別の色を割り当てて表示することで3群以上の比較表示も可能である。

【0019】

なお、本実施例では、赤、青などの色を用いているがグレイスケールや、他の模

様を用いことも可能である。また、一塩基多型データで説明しているが、マイクロサテライトなどのデータであってもペアサイズの分割表を作成し、その独立性の検定を行い、カイ二乗値や、そこから求められるP値を用いて同様に連鎖不平衡の代わりにして、そのまま画像として表示することも可能である。

【0020】

また、文献K. Shimo-onoda et al : Akaike's information criterion for a measure of linkage disequilibrium, Journal of Human Genetics, Vol. 47 Issue 12 (2002) p 649-655に示すようにAICの独立もモデルと、従属モデルを定義してその差をとった連鎖不平衡値を利用するのも可能である。カイ二乗値やAICによる連鎖不平衡の値を用いる場合にはその値の範囲が0以上の広い範囲に及ぶため、連鎖不平衡値を実際に求めた値の最大値を探査し、その最大値に対して各色をマッピングすることで、同様に視覚的に分かりやすいグラフィック表示を行うことができる。

【0021】

【発明の効果】 以上、説明したように本発明の遺伝子連鎖不平衡の表示方法によれば、グラフィカルに連鎖不平衡の値を表示するため、複数の群の相違を容易に比較評価ができるという優れた効果を持つ。

【図面の簡単な説明】

【図1】

本発明の一実施例を説明するための全体構成図

【図2】

入力データとケース・コントロール群の連鎖不平衡値を算出した例

【図3】

連鎖不平衡値を色で表現して表示する例

【図4】

色別に表示するための処理フロー例

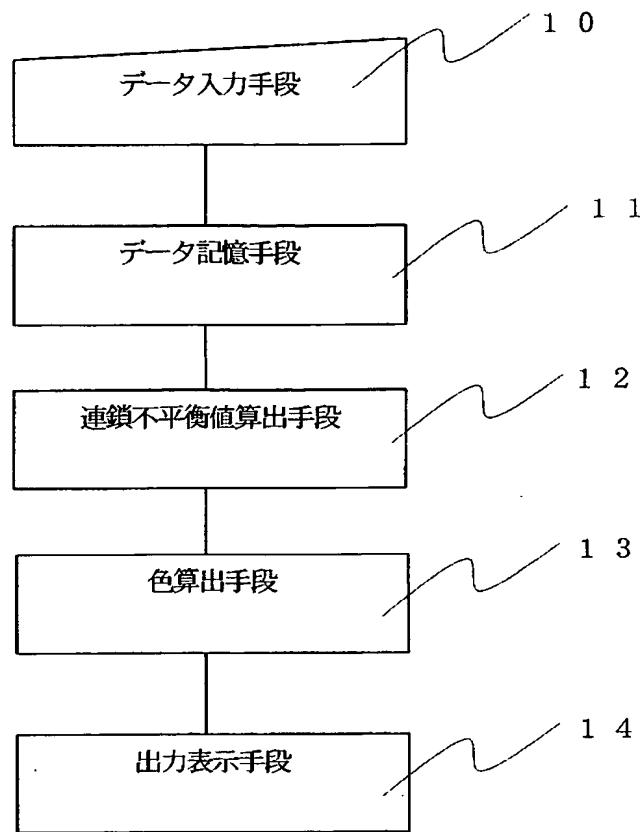
【符号の説明】

- 1 0 データ入力手段
- 1 1 データ記憶手段
- 1 2 連鎖不平衡値算出手段
- 1 3 色算出手段
- 1 4 出力表示手段
- 3 1 ~ 3 6 各セルの色表示
- 4 0 ~ 4 5 色表示のための処理フロー

【書類名】

図面

【図1】



【図2】

入力データ

	SNP1	SNP2	SNP3	SNP4	群
ID1	1	3	1	3	0
ID2	1	2	1	2	0
ID3	1	2	1	3	1
:	:	:	:	:	:

ケースデータ群の連鎖不平衡値

	SNP1	SNP2	SNP3	SNP4
SNP1		0.1	0.9	0.1
SNP2			0.1	0.9
SNP3				0.1
SNP4				

コントロールデータ群の連鎖不平衡値

	SNP1	SNP2	SNP3	SNP4
SNP1		0.1	0.9	0.0
SNP2			0.1	1.0
SNP3				0.0
SNP4				

【図3】

3 1 3 2

(A) 2色混合

	SNP1	SNP2	SNP3	SNP4
SNP1		■■■	■■■	■■■■
SNP2			■■■	■■■■
SNP3				■■■■
SNP4				

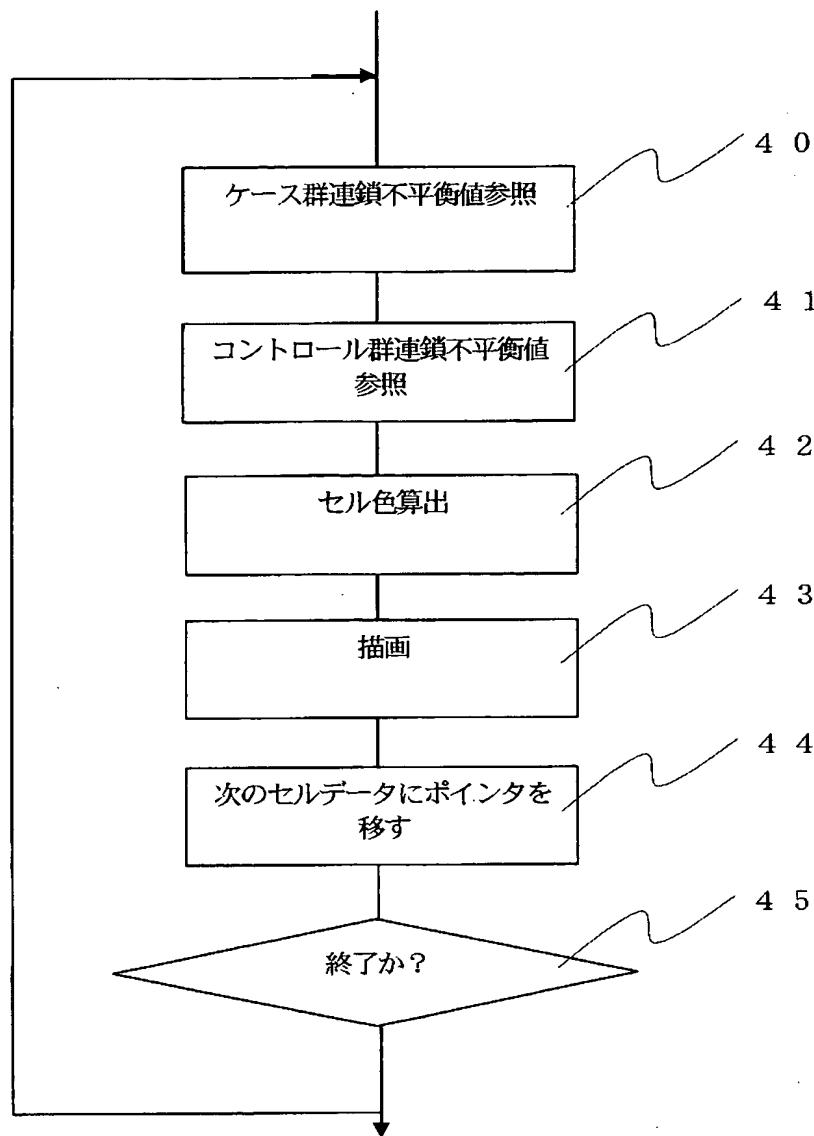
3 3
3 4

(B) 差分表示

	SNP1	SNP2	SNP3	SNP4
SNP1				■■■■
SNP2				■■■■
SNP3				■■■■
SNP4				

3 5
3 6

【図4】



【書類名】 要約書

【要約】

【課題】 罹患患者と健常者などの2つ以上のグループから取得された遺伝子多様性データからペアワイズに求めた連鎖不平衡値の差を分かりやすく表示する。

【解決手段】 2つ以上の遺伝子座位における遺伝子多様性データを入力して連鎖不平衡解析を行う場合に、疾患などのケース群と対象となるコントロール群に分けて、連鎖不平衡を求める。さらにそれぞれのグループの連鎖不平衡値を色の濃淡に対応させる。さらに2つのグループの色を混色して表示することで、両者に差のある遺伝子多型を効率よく特定できるようにする。また、差だけに着目する場合にはあらかじめケース群とコントロール群の連鎖不平衡値の差を求めておき、その値に応じて色を変えて表示させることで、両者の比較を視覚的に行えるようにする。

【選択図】 図1

特願 2003-048216

出願人履歴情報

識別番号 [597101155]

1. 変更年月日 2002年 2月 7日

[変更理由] 住所変更

住 所 千葉県茂原市茂原643番地
氏 名 株式会社ダイナコム